APEX DIGITAL SYSTEMS
1010 Wayne Avenue, Suite 800, Silver Spring MD 20910
301-588-9767 fax: 301-588-9766
www.apexdigitalsystems.com

Analysis Report: XML Use at NWS River Forecast Centers

Summary Findings of Current XML Use at the 13 River Forecast Centers (RFCs)

Stephanie Liu-Barnes, Director of Consulting and Design
Andrey Ignatov, Data Architect

December 22, 2005

Summary

The National Weather Service's Office of Hydrologic Development and Apex Digital Systems have worked together for over a year on establishing and growing the Hydrologic XML Consortium (HydroXC) for the purpose of creating a common, self-documenting method for exchanging hydrologic data among the many organizations that use such data both operationally and for research.

As part of this effort, Apex was charged with the task to survey all 13 River Forecast Centers (RFCs) regarding their knowledge and use of XML in any form. The Apex team surveyed 11 of the 13 RFCs with a specific protocol of questions (attached as Appendix B) over the course of several months.

This document contains a summary and analysis of the various conversations and the resulting information.

Background: Hydrology XML Consortium (HydroXC)

The Hydrology XML Consortium is a loose affiliation of organizations working in the hydrologic field, or using hydrologic data on a daily basis. Currently, the Consortium includes representatives from the National Weather Service and several River Forecast Centers, the Army Corps of Engineers, the US Geological Survey, several universities, and private sector businesses.

The main objective of the consortium is to open channels of communication among the various hydrologic practitioners around the topic of standardizing data exchange capabilities. The Extensible Markup Language (XML) has been adapted in many sector-specific areas (finance, insurance, medical research), and the efforts of the HydroXC has already resulted in an initial schema for open, self-documenting data exchange capabilities based on XML.

The Survey Process

The Office of Hydrologic Development (OHD) identified one contact person at each of the 13 RFCs and Apex created a standard list of questions for the XML survey (see Appendix B). The questions focused on extensible markup language (XML), how it is or might be used, and general awareness of HydroXC. Apex contacted each RFC and asked to schedule a short phone conversation for the survey.

Two Apex team members conducted the survey over the phone with one or more participants at each of the responding RFC.

Apex interviewed 11 of the 13 RFCs. APRFC in Anchorage, Alaska declined the interview, and the Apex team was unable to make contact with OHRFC in Wilmington, Ohio.

Findings

The conversations with the individual RFCs consisted of four discussion areas:

- Awareness of the Hydrologic XML Consortium,
- Current use or knowledge of XML within the RFC,
- The need for XML-based modeling in hydrology, and
- Future thinking, addressing concerns and opportunities for XML in hydrology.

The majority of the RFCs had heard about HydroXC, mostly from visiting the website or talking briefly with someone at headquarters (OHD), but this awareness was mostly peripheral and did not go into understanding of the Consortium's current progress or tasks.

The majority of respondents also understood the basics of XML, but had not implemented specific projects using XML at their RFC, or even conducted significant discussions about XML within their office. Three of the RFCs had begun to delve deeply into the practical uses of XML.

Most RFCs seemed appreciative of being included in this phase of the HydroXC discovery, but did not express interest in becoming active members of HydroXC. Reasons for this included lack of staffing or resources for the additional work. Two RFCs ruled out participation altogether at this time, stating they would like to remain on e-mail lists for updates. Three RFCs expressed enthusiasm about participating in the development of a standard hydrologic XML schema.

Detailed findings for each discussion area follow.

1: Awareness of the HydroXC

While OHD, Apex and the HydroXC members have been working closely together to develop a suitable XML schema for hydrology, the majority of the RFCs have, so far, not been direct contributors to the project. During each of the interviews, the RFC participants requested a detailed update on the background and activities of the Consortium to date.

Overall, the RFCs appear to have a general awareness of the HydroXC. Only two of the eleven RFCs reported having no familiarity with the Consortium at all (MBRFC and SERFC). The other RFCs had at least heard of it, but were not familiar with its tasks and goals. Most of the interviewees had visited the HydroXC website and a few had heard of HydroXC during general conversations with staff at OHD. Many said they found the web site lacking in specific details, but were able to glean a general concept of HydroXC's purpose.

The Apex team also asked each interviewee to describe their expectations of the HydroXC. Most expected a standard schema that would be mandated by OHD and used in AWIPS data exchanges, in some capacity. Several RFCs had questions about when they might be expected to integrate XML into their daily operations and how they would be trained to use the new technology.

2: Current use of XML and other data exchanges

Next the Apex team explored the RFCs' current use or study of XML.

Four RFCs said they were currently using XML - CBRFC, CNRFC, NCRFC, and NWRFC, as follows:

- CBRFC (Colorado Basin RFC) is using XML for Water Supply Forecasts, mainly used in the western United States. The Water Supply Forecasts have been developed in collaboration with NWRFC and have included discussions with CNRFC, MBRFC and WGRFC. These forecasts predict how much water will run off of mountains and into reservoirs, with the goal of producing one web page to represent the water supply for all western RFC areas. According to Development Operations Hydrologist Steven Shumate, the project is close to completion, with some fine-tuning of graphic displays outstanding. CBRFC also uses XML to represent some event, time series, and climate data from their FASTETC database.
- CNRFC (California Nevada RFC) has worked closely with CBRFC to develop the Water Supply Forecasts for the western United States. Otherwise, the staff there has not yet identified further areas for XML development.
- NCRFC (North Central RFC) is actively investigating XML use in hydrology, driven mostly by their Information Technology Officer, Dan Pokorny. Dan worked for the RFC years ago and left for private sector work. He returned earlier this year to the RFC with XML knowledge and experience. Together with Development and Operations Hydrologist, John Halquist, Dan is involved in several initiatives that incorporate XML. One active application produces web graphics from an ensemble forecasting tool and uses XML to produce hydrographs. Dan is also involved in a national effort among RFCs, and in association with the U.S. Geological Survey (USGS), to develop a method for exchanging rating curve data from the USGS to the RFCs in a consistent way. XML is being used in parts of that project as well.
- NWRFC (Northwest RFC) has been working on a hydro-forecast product using XML, but is waiting for a standard from HydroXC before making the product available to the public. They are also collaborating with CBRFC on the western region water supply forecast web page. Additionally, they run four RSS (Really Simple Syndication) feeds from their web site. The RSS feed is a family of XML file formats used primarily among news organizations and does not specifically support hydrologic data. However, NWRFC has defined a simple way to represent hydrologic data in the RSS format, thus enabling the feeds of hydrologic data.

Most of the other RFCs have no specific projects underway involving XML, but did describe data and data exchanges they use that would benefit from integration with XML, as follows:

- Rating curve data that is exchanged with the USGS on a regular basis
- A variety of data exchanges with the U.S. Army Corps of Engineers
- Ensemble forecast data
- Output from the FLDWAV application
- Observation and forecast data that is now encoded into SHEF (Standard Hydrometeorological Exchange Format)
- Any data that is exchanged with Weather Forecast Offices (WFOs)
- Data exchanges with local authorities, such as the Tennessee Valley Authority (TVA), the Regional Climate Center at Cornell University, or the Alabama Power Company

Several respondents also identified specific data formats for which they would like to see conversion utilities:

- XMRG data
- LDM (Local Data Manager) backup data transmissions
- GRIB (Gridded Binary) data
- GIS shapefile spatial data
- NetCDF (network Common Data Form)

Finally, interviewees also indicated interest in integrating HydroXC-compliant XML with:

- AHPS (Advanced Hydrologic Prediction Service) HydroGen. HydroGen is suite of software programs that prepares XML files and hydrograph graphics for the web.
- The standard RSS format to better support hydrology

Of the RFCs currently using XML, all have staff members on site with XML capabilities of varying degrees. Also, MARFC and ABRFC both reported having staff who were experimenting with XML, although they have not defined specific XML projects. The majority of RFCs did say they have people on staff who would be interested in learning about the technology.

3: Need for XML in hydrology

When asked if they thought it was important to explore XML use with hydrology, most interviewees agreed that it was important to modernize the current hydrologic data standard, SHEF. However, a few RFCs, including LMRFC and WGRFC, indicated that SHEF meets their current data needs, internally and in exchanges with other organizations, and did not see immediate benefit to changing standards at this point in time. The National Weather Service has been using SHEF for many years, with the result that many data exchanges, while complex, operate very reliably.

When questioned further, however, all RFCs, except LMRFC, described scenarios or situations in which exchanging data in the SHEF format has been problematic. Current concerns around SHEF included the fact that many of the long-term RFC staff members are beginning to retire. Many of the new hydrologists and meteorologists coming into the RFCs do not know SHEF and find that the format is difficult to grasp fully.

Similarly, many of the outside agencies and groups also have someone in-house that can maintain the data exchanges with the RFC. However, when that person leaves, it is very difficult to train other staff members on the use of the SHEF format. The interviewees indicated that using XML with its self-documenting data modeling format should significantly reduce training requirements for new staff members.

At this point, many RFCs do not believe that they have sufficient experience with XML to know if XML is indeed the most appropriate modeling environment for modernizing forecasting software. However, they do appear to generally support the approach and are willing to learn the standards, as they become available.

One of the concerns the Apex team heard in several interviews about XML is the tendency of XML data files to become very large. While the self-describing schema makes data files easier to use, the self-describing tag structure adds significant volume to the data set. MARFC expressed the concern that the self-describing structure in XML data files would result in lower readability compared to the compact SHEF format.

4: Future thinking and concerns

The RFCs surveyed by the Apex team expressed several useful concerns about adopting XML in hydrology.

The difficulty of the transition to XML was a recurring concern, with several RFCs specifically identifying customer education. While it has been difficult to educate and train data consumers and providers to encode to and decode from SHEF, most RFCs operate their SHEF-based data exchanges relatively smoothly today. The idea of having to “start from scratch” and retrain these groups to use a different format seemed overwhelming. Most mentioned that, first and foremost, RFC staff will have to be educated about XML and the hydrologic standard adopted by OHD. Secondly, all data consumers and suppliers would have to be educated about the new standards and current practices updated to support it.

In order to make a smooth transition to the hydrologic XML standard, several RFCs asked for standard data decoders and encoders, to be used with today’s most highly used formats, esp. SHEF. They also asked for practical examples and code libraries that are easy to reference. We also heard the need for continued progress toward these goals, so that the current progress and enthusiasm did not lose momentum.

Two RFCs, in particular, expressed concerns that the hydrologic XML modeling effort was moving too slowly and/or was not fully a cooperative effort among organizations. ABRFC mentioned concern that the project has already been active for over a year, without producing prototypes and actual system examples to show its usefulness. NWRFC also said the project was moving slower than they would like and asked to see more practical applications earlier in the process. They were excited about the overall initiative, but felt the real value would be in showing how a few real applications might use the HydroXC structure to exchange data with each other.

Many interviewees also are looking for more active participation by other agencies in showing proofs of concept and felt the only progress to date required “paid work” by Apex. They understood some of the complications of organizing a cross-organization standard, but wanted to see continued interest and involvement from organizations outside of the NWS.

Out of all the interviewed RFCs, all said they wanted to remain on the e-mail list to be informed of further developments to do with HydroXC and XML use at the RFCs. As a result of the survey conversation and subsequent discussions, NCRFC formally joined the Hydrologic XML Consortium as an active participant and collaborator.

Conclusion

Reaching out to all of the RFCs was an important step in the development of the Hydrologic XML draft schema. By ensuring all of the RFCs have an understanding of the progress and vision and enabling them to participate and contribute, OHD will be able to work more effectively with the RFCs in defining and implementing XML-based modeling into daily operations.

Several key themes emerged from the survey activity, which OHD should consider in depth for further action:

- The HydroXC work should be an on-going effort: After more than a year of groundwork, the RFCs, and many of the other organizations, are looking to the HydroXC activities as a thought-leader regarding open data exchange. As a result, the Consortium participants notice, and comment on, the fact that the activities are intermittent based on Apex's contracts with OHD.

  Recommendation: Ideally, the work of the Consortium should continue throughout the year to meet the growing interest in and demand for its leadership.

- Provide education & training: Interest in implementing an open XML-based schema for data exchanges is strong among the RFCs. At the same time, they recognize the need for training their staff to make effective use of emerging new technologies.

  Recommendation: The Consortium should plan to provide guidelines for such training and possibly provide some of this training in the context of the HydroXC schema.

- Implement and distribute software tools: RFCs and other collaborators will need basic software tools to read and write HydroXC XML data. Specifically, RFC interest in a code base for converting between SHEF and XML appears to be high.

  Recommendation: The Consortium should implement and distribute open source code and utilities for converting data between various formats and HydroXC XML. The Consortium should also create and distribute sample applications and datasets that developers and hydrologists can use to test and evaluate the code and utilities.

- Continue to engage the community: The RFC interviewees stated strongly that continued and broadening engagement with organizations outside of the NWS is important to them. The inclusion of USGS, US Dept. of Agriculture, Army Corps of Engineers, Drexel University's Michael Piasecki, and several commercial companies differentiates the Consortium in many positive ways.

  Recommendation: The Consortium should continue to foster stronger ties with a wide range of organizations both inside and outside the federal government.

- Strengthen the visibility of OHD leadership: RFC interviewees expressed the desire to see OHD's leadership connect itself visibly and strongly to the project, to avoid the Consortium appearing as a contractor-centric effort.

  Recommendation: OHD leaders should continue to participate in, and contribute to, the work of the Consortium. Further, any efforts by OHD to implement the fruits of the Consortium's work in OHD systems will show tangible results that many RFCs are hoping to see.

APPENDIX A: RFC participants of the XML survey

ABRFC (Arkansas-Red Basin RFC)
Location: Tulsa, OK
Participants: Ken Pavelle
Hydrologist / Webmaster
Ken.Pavelle@noaa.gov
918-832-4110

CBRFC (Colorado Basin RFC)
Location: Salt Lake City, UT
Participants: Steven Shumate
Development Operational Hydrologist
Steven.Shumate@noaa.gov
801-524-5130

CNRFC (California Nevada RFC)
Location: Sacramento, CA
Participants: Robert Hartman
Hydrologist in Charge
Robert.Hartman@noaa.gov
916-979-3056 x322

LMRFC (Lower Mississippi RFC)
Location: Slidell, LA
Participants: Eric Jones
Senior Hydrologist
Eric.Jones@noaa.gov
985-641-4343 x238

MARFC (Middle Atlantic RFC)
Location: State College, PA
Participants: Joe Ostrowski
Development and Operational Hydrologist
Joe.Ostrowski@noaa.gov
814-231-2403

MBRFC (Missouri Basin RFC)
Location: Pleasant Hill, MO
Participants: Thomas Gurss
Development and Operational Hydrologist
Tom.Gurss@noaa.gov
816-540-5151

John Lague
Senior HAS Forecaster
John.Lague@noaa.gov
816-540-5151 x743

NCRFC (North Central RFC)
Location: Chanhassen, MN
Participants:

John Halquist
Development and Operations Hydrologist
John.Halquist@noaa.gov
952-361-6650, ext. 346

Dan Pokorny
Information Technology Officer
Dan.Pokorny@noaa.gov
952-361-6664, ext. 384

NERFC (Northeast RFC)
Location: Taunton, MA
Participants: Robert Shedd
Development & Operations Hydrologist
Robert.Shedd@noaa.gov
508-824-5116 x233

NWRFC (Northwest RFC)
Location: Portland, OR
Participants:

Harold Optiz
Hydrologist In Charge
Harold.Opitz@noaa.gov
503-326-7401

Don Laurine
Development and Operations Hydrologist
Donald.Laurine@noaa.gov

SERFC (Southeast RFC)
Location: Peachtree City, GA
Participants:

Jonathan Atwell
Senior Hydrologist
Jonathan.Atwell@noaa.gov
770-486-0028

Brad Gimmestad
Development and Operations Hydrologist

Judy Bradbury
Senior Hydrometeorological Analysis and Support (HAS) Forecaster
Hydrologist

WGRFC (West Gulf RFC)
Location: Forth Worth, TX
Participants: Bob Corby
Development and Operations Hydrologist
Robert.Corby@noaa.gov
817-831-3289

APPENDIX B: RFC Survey Questionnaire

The following questionnaire guided conversations with each RFC during the XML survey phone calls.